



IFWO

RAW SEQUENCE LISTING

DATE: 07/30/2004

PATENT APPLICATION: US/10/796,669A

TIME: 11:26:55

Input Set : A:\741179.txt

Output Set: N:\CRF4\07302004\J796669A.raw

ENTERED

```

3 <110> APPLICANT: Wang, Jin-Town
4   Lin, Tzu-Lung
6 <120> TITLE OF INVENTION: A TYPE II RESTRICTION ENDONUCLEASE AND APPLICATION THEREOF
8 <130> FILE REFERENCE: P/741-179
10 <140> CURRENT APPLICATION NUMBER: 10/796,669A
11 <141> CURRENT FILING DATE: 2004-03-09
13 <160> NUMBER OF SEQ ID NOS: 5
15 <170> SOFTWARE: PatentIn version 3.2
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 5
19 <212> TYPE: DNA
20 <213> ORGANISM: Helicobacter pylori
22 <400> SEQUENCE: 1
23 ccatc
26 <210> SEQ ID NO: 2
27 <211> LENGTH: 1617
28 <212> TYPE: DNA
29 <213> ORGANISM: Helicobacter pylori
32 <220> FEATURE:
33 <221> NAME/KEY: CDS
34 <222> LOCATION: (1)..(1617)
36 <400> SEQUENCE: 2
37 atg act aaa aaa ccg gca cga aaa att tta agc ttt tca acc acc atg
38 Met Thr Lys Lys Pro Ala Arg Lys Ile Leu Ser Phe Ser Thr Thr Met
39 1 5 10 15
41 cga aac cct aaa aga ata gga caa ttt tta gct gtt tta gga aag ttt
42 Arg Asn Pro Lys Arg Ile Gly Gln Phe Leu Ala Val Leu Gly Lys Phe
43 20 25 30
45 gaa aat caa atc ctt aaa tct tca ata atc atg caa att atc aaa tcc
46 Glu Asn Gln Ile Leu Lys Ser Ser Ile Ile Met Gln Ile Ile Lys Ser
47 35 40 45
49 gtt ttg gct cat agg ctt tat aga cct act tct ctc aat caa aat aaa
50 Val Leu Ala His Arg Leu Tyr Arg Pro Thr Ser Leu Asn Gln Asn Lys
51 50 55 60
53 gaa ttg aaa gaa aaa ttt gac tcc aat gaa tat gtc ttt agc gat gaa
54 Glu Leu Lys Glu Lys Phe Asp Ser Asn Glu Tyr Val Phe Ser Asp Glu
55 65 70 75 80
57 gag tta gaa cgc att ata gaa ata tcc cca caa aat cat aaa gaa atg
58 Glu Leu Glu Arg Ile Ile Glu Ile Ser Pro Gln Asn His Lys Glu Met
59 85 90 95
61 ggc ttt gag cat gga tgg gaa agc cgg ttt gac act tgg tat aag ctt
62 Gly Phe Glu His Gly Trp Glu Ser Arg Phe Asp Thr Trp Tyr Lys Leu
63 100 105 110

```

RAW SEQUENCE LISTING

DATE: 07/30/2004

PATENT APPLICATION: US/10/796,669A

TIME: 11:26:55

Input Set : A:\741179.txt

Output Set: N:\CRF4\07302004\J796669A.raw

65	atg	tgt	gag	ttt	ggt	ttt	tgc	tac	tat	gca	aaa	tat	gag	aaa	ata	ctc	384
66	Met	Cys	Glu	Phe	Gly	Phe	Cys	Tyr	Tyr	Ala	Lys	Tyr	Glu	Lys	Ile	Leu	
67			115				120						125				
69	atc	agc	gat	agc	gct	aag	atg	ctt	att	ctt	gct	tat	tac	aat	aaa	gaa	432
70	Ile	Ser	Asp	Ser	Ala	Lys	Met	Leu	Ile	Leu	Ala	Tyr	Tyr	Asn	Lys	Glu	
71			130				135						140				
73	aac	gat	gct	ttt	aaa	gaa	agc	ggt	gat	gaa	agc	gta	ggt	ggg	gct	ata	480
74	Asn	Asp	Ala	Phe	Lys	Glu	Ser	Val	Asp	Glu	Ser	Val	Val	Gly	Ala	Ile	
75	145					150						155				160	
77	ttt	tta	aac	gct	ctg	tct	aaa	tat	gaa	gta	gga	aac	cct	tac	aaa	aag	528
78	Phe	Leu	Asn	Ala	Leu	Ser	Lys	Tyr	Glu	Val	Gly	Asn	Pro	Tyr	Lys	Lys	
79				165						170				175			
81	aat	tta	aac	cat	aac	aac	cct	ttc	aaa	cta	ttg	ctc	tcg	ctt	tta	aaa	576
82	Asn	Leu	Asn	His	Asn	Asn	Pro	Phe	Lys	Leu	Leu	Leu	Ser	Leu	Leu	Lys	
83				180						185				190			
85	cga	ctc	aaa	aat	gcc	cat	cta	acc	ccc	cta	tct	gtc	aaa	gaa	atc	cct	624
86	Arg	Leu	Lys	Asn	Ala	His	Leu	Thr	Pro	Leu	Ser	Val	Lys	Glu	Ile	Pro	
87			195					200					205				
89	att	tta	ctt	tgt	tgg	aaa	gac	gat	aac	gct	aat	ggg	ctt	tat	gac	tac	672
90	Ile	Leu	Leu	Cys	Trp	Lys	Asp	Asp	Asn	Ala	Asn	Gly	Leu	Tyr	Asp	Tyr	
91			210				215					220					
93	att	att	cgt	tta	aga	caa	gaa	atc	ggt	act	atc	aat	aaa	aca	gaa	ttc	720
94	Ile	Ile	Arg	Leu	Arg	Gln	Glu	Ile	Val	Thr	Ile	Asn	Lys	Thr	Glu	Phe	
95	225					230					235					240	
97	agc	tac	tca	gat	gaa	ttt	atc	tat	gaa	aaa	tgc	cta	aaa	ctt	tta	gaa	768
98	Ser	Tyr	Ser	Asp	Glu	Phe	Ile	Tyr	Glu	Lys	Cys	Leu	Lys	Leu	Leu	Glu	
99				245						250				255			
101	agt	ggt	aat	aaa	aca	cga	ttt	aaa	atg	agc	caa	atc	act	aac	gaa	gcc	816
102	Ser	Val	Asn	Lys	Thr	Arg	Phe	Lys	Met	Ser	Gln	Ile	Thr	Asn	Glu	Ala	
103				260						265				270			
105	ggt	gat	gaa	tac	att	aga	aaa	atg	cgt	att	aca	gga	ctt	att	tca	ttg	864
106	Val	Asp	Glu	Tyr	Ile	Arg	Lys	Met	Arg	Ile	Thr	Gly	Leu	Ile	Ser	Leu	
107			275					280					285				
109	cgt	ggt	aat	ggt	agg	ttt	att	gat	att	aat	act	aat	gaa	aat	aat	aaa	912
110	Arg	Gly	Asn	Gly	Arg	Phe	Ile	Asp	Ile	Asn	Thr	Asn	Glu	Asn	Asn	Lys	
111			290				295					300					
113	ata	gat	tac	att	tta	caa	acc	cat	aag	gct	ttt	aaa	ggg	gat	tat	tta	960
114	Ile	Asp	Tyr	Ile	Leu	Gln	Thr	His	Lys	Ala	Phe	Lys	Gly	Asp	Tyr	Leu	
115	305					310					315					320	
117	aac	gac	act	caa	gct	aac	aaa	ctc	gcc	ttt	ttt	aac	tac	atg	gcg	atc	1008
118	Asn	Asp	Thr	Gln	Ala	Asn	Lys	Leu	Ala	Phe	Phe	Asn	Tyr	Met	Ala	Ile	
119				325						330				335			
121	gtg	gat	agc	ttt	ctt	ggt	agt	ggt	act	cca	atc	agc	gct	aat	gag	agc	1056
122	Val	Asp	Ser	Phe	Leu	Val	Ser	Val	Thr	Pro	Ile	Ser	Ala	Asn	Glu	Ser	
123				340						345				350			
125	ggt	aaa	tca	agc	aaa	ttg	aat	gaa	cta	gca	aac	act	tat	act	aaa	gat	1104
126	Val	Lys	Ser	Ser	Lys	Leu	Asn	Glu	Leu	Ala	Asn	Thr	Tyr	Thr	Lys	Asp	
127			355						360				365				
129	ttt	atc	aag	caa	gaa	tta	ctc	att	act	tgt	aac	aag	caa	gaa	tca	aaa	1152

RAW SEQUENCE LISTING

DATE: 07/30/2004

PATENT APPLICATION: US/10/796,669A

TIME: 11:26:55

Input Set : A:\741179.txt

Output Set: N:\CRF4\07302004\J796669A.raw

```

130 Phe Ile Lys Gln Glu Leu Leu Ile Thr Cys Asn Lys Gln Glu Ser Lys
131      370      375      380
133 gat agt ttt tta aga ctc att gat aaa cct tta cgc tta gaa ttt tta      1200
134 Asp Ser Phe Leu Arg Leu Ile Asp Lys Pro Leu Arg Leu Glu Phe Leu
135 385      390      395      400
137 agc gct att ttc ttg aaa caa cat ttt gaa aat tta agc gtg ata ccc      1248
138 Ser Ala Ile Phe Leu Lys Gln His Phe Glu Asn Leu Ser Val Ile Pro
139      405      410      415
141 aat tat aaa agc gat gat gaa ggc ttg ccc gta tac aca gca agc ggt      1296
142 Asn Tyr Lys Ser Asp Asp Glu Gly Leu Pro Val Tyr Thr Ala Ser Gly
143      420      425      430
145 aat aaa cct gat att gta gct atg gac aca aaa gcc caa agt tat ata      1344
146 Asn Lys Pro Asp Ile Val Ala Met Asp Thr Lys Ala Gln Ser Tyr Ile
147      435      440      445
149 gaa gtg agc ttg att aga gac aga agt caa agt acc ttg gaa atg ata      1392
150 Glu Val Ser Leu Ile Arg Asp Arg Ser Gln Ser Thr Leu Glu Met Ile
151      450      455      460
153 cct att gcc aga cat tta aaa gaa ttg att aaa aat agc acc gat att      1440
154 Pro Ile Ala Arg His Leu Lys Glu Leu Ile Lys Asn Ser Thr Asp Ile
155 465      470      475      480
157 aga gaa aaa ttt agt gtt ttt gta gct cca aat atc cat gat gat gcc      1488
158 Arg Glu Lys Phe Ser Val Phe Val Ala Pro Asn Ile His Asp Asp Ala
159      485      490      495
161 aaa gaa tat gcg gaa ttt gcc caa ttc aaa gac aat att aat ata tgt      1536
162 Lys Glu Tyr Ala Glu Phe Ala Gln Phe Lys Asp Asn Ile Asn Ile Cys
163      500      505      510
165 tgt tat gct att aat gat ttt atc aaa aaa gta gaa aac agc ata gaa      1584
166 Cys Tyr Ala Ile Asn Asp Phe Ile Lys Lys Val Glu Asn Ser Ile Glu
167      515      520      525
169 tgg tta cag atc aat gac cat ttg aaa gct taa      1617
170 Trp Leu Gln Ile Asn Asp His Leu Lys Ala
171      530      535
174 <210> SEQ ID NO: 3
175 <211> LENGTH: 538
176 <212> TYPE: PRT
177 <213> ORGANISM: Helicobacter pylori
179 <400> SEQUENCE: 3
181 Met Thr Lys Lys Pro Ala Arg Lys Ile Leu Ser Phe Ser Thr Thr Met
182 1      5      10      15
185 Arg Asn Pro Lys Arg Ile Gly Gln Phe Leu Ala Val Leu Gly Lys Phe
186      20      25      30
189 Glu Asn Gln Ile Leu Lys Ser Ser Ile Ile Met Gln Ile Ile Lys Ser
190      35      40      45
193 Val Leu Ala His Arg Leu Tyr Arg Pro Thr Ser Leu Asn Gln Asn Lys
194      50      55      60
197 Glu Leu Lys Glu Lys Phe Asp Ser Asn Glu Tyr Val Phe Ser Asp Glu
198 65      70      75      80
201 Glu Leu Glu Arg Ile Ile Glu Ile Ser Pro Gln Asn His Lys Glu Met
202      85      90      95

```

RAW SEQUENCE LISTING

DATE: 07/30/2004

PATENT APPLICATION: US/10/796,669A

TIME: 11:26:55

Input Set : A:\741179.txt

Output Set: N:\CRF4\07302004\J796669A.raw

```

205 Gly Phe Glu His Gly Trp Glu Ser Arg Phe Asp Thr Trp Tyr Lys Leu
206          100          105          110
209 Met Cys Glu Phe Gly Phe Cys Tyr Tyr Ala Lys Tyr Glu Lys Ile Leu
210          115          120          125
213 Ile Ser Asp Ser Ala Lys Met Leu Ile Leu Ala Tyr Tyr Asn Lys Glu
214          130          135          140
217 Asn Asp Ala Phe Lys Glu Ser Val Asp Glu Ser Val Val Gly Ala Ile
218 145          150          155          160
221 Phe Leu Asn Ala Leu Ser Lys Tyr Glu Val Gly Asn Pro Tyr Lys Lys
222          165          170          175
225 Asn Leu Asn His Asn Asn Pro Phe Lys Leu Leu Leu Ser Leu Leu Lys
226          180          185          190
229 Arg Leu Lys Asn Ala His Leu Thr Pro Leu Ser Val Lys Glu Ile Pro
230          195          200          205
233 Ile Leu Leu Cys Trp Lys Asp Asp Asn Ala Asn Gly Leu Tyr Asp Tyr
234          210          215          220
237 Ile Ile Arg Leu Arg Gln Glu Ile Val Thr Ile Asn Lys Thr Glu Phe
238 225          230          235          240
241 Ser Tyr Ser Asp Glu Phe Ile Tyr Glu Lys Cys Leu Lys Leu Leu Glu
242          245          250          255
245 Ser Val Asn Lys Thr Arg Phe Lys Met Ser Gln Ile Thr Asn Glu Ala
246          260          265          270
249 Val Asp Glu Tyr Ile Arg Lys Met Arg Ile Thr Gly Leu Ile Ser Leu
250          275          280          285
253 Arg Gly Asn Gly Arg Phe Ile Asp Ile Asn Thr Asn Glu Asn Asn Lys
254          290          295          300
257 Ile Asp Tyr Ile Leu Gln Thr His Lys Ala Phe Lys Gly Asp Tyr Leu
258 305          310          315          320
261 Asn Asp Thr Gln Ala Asn Lys Leu Ala Phe Phe Asn Tyr Met Ala Ile
262          325          330          335
265 Val Asp Ser Phe Leu Val Ser Val Thr Pro Ile Ser Ala Asn Glu Ser
266          340          345          350
269 Val Lys Ser Ser Lys Leu Asn Glu Leu Ala Asn Thr Tyr Thr Lys Asp
270          355          360          365
273 Phe Ile Lys Gln Glu Leu Leu Ile Thr Cys Asn Lys Gln Glu Ser Lys
274          370          375          380
277 Asp Ser Phe Leu Arg Leu Ile Asp Lys Pro Leu Arg Leu Glu Phe Leu
278 385          390          395          400
281 Ser Ala Ile Phe Leu Lys Gln His Phe Glu Asn Leu Ser Val Ile Pro
282          405          410          415
285 Asn Tyr Lys Ser Asp Asp Glu Gly Leu Pro Val Tyr Thr Ala Ser Gly
286          420          425          430
289 Asn Lys Pro Asp Ile Val Ala Met Asp Thr Lys Ala Gln Ser Tyr Ile
290          435          440          445
293 Glu Val Ser Leu Ile Arg Asp Arg Ser Gln Ser Thr Leu Glu Met Ile
294          450          455          460
297 Pro Ile Ala Arg His Leu Lys Glu Leu Ile Lys Asn Ser Thr Asp Ile
298 465          470          475          480
301 Arg Glu Lys Phe Ser Val Phe Val Ala Pro Asn Ile His Asp Asp Ala

```

RAW SEQUENCE LISTING

DATE: 07/30/2004

PATENT APPLICATION: US/10/796,669A

TIME: 11:26:55

Input Set : A:\741179.txt

Output Set: N:\CRF4\07302004\J796669A.raw

```

302          485          490          495
305 Lys Glu Tyr Ala Glu Phe Ala Gln Phe Lys Asp Asn Ile Asn Ile Cys
306          500          505          510
309 Cys Tyr Ala Ile Asn Asp Phe Ile Lys Lys Val Glu Asn Ser Ile Glu
310          515          520          525
313 Trp Leu Gln Ile Asn Asp His Leu Lys Ala
314          530          535
317 <210> SEQ ID NO: 4
318 <211> LENGTH: 780
319 <212> TYPE: DNA
320 <213> ORGANISM: Helicobacter pylori
323 <220> FEATURE:
324 <221> NAME/KEY: CDS
325 <222> LOCATION: (1)..(780)
327 <400> SEQUENCE: 4
328 atg ggg caa gac gct gat ttt aaa gcg ctt gaa gaa ctg aaa gaa tac      48
329 Met Gly Gln Asp Ala Asp Phe Lys Ala Leu Glu Glu Leu Lys Glu Tyr
330 1          5          10          15
332 ttt aat caa gct tta aag cta gaa gaa aat tat ttt agc caa cat ttt      96
333 Phe Asn Gln Ala Leu Lys Leu Glu Glu Asn Tyr Phe Ser Gln His Phe
334          20          25          30
336 agc aac aag ttt ttc agc tat aaa gat tgt gtc aaa atc ggt agc att      144
337 Ser Asn Lys Phe Phe Ser Tyr Lys Asp Cys Val Lys Ile Gly Ser Ile
338          35          40          45
340 aga gag cat ata gaa agc tta aac tta gat aaa tta aat aaa gat att      192
341 Arg Glu His Ile Glu Ser Leu Asn Leu Asp Lys Leu Asn Lys Asp Ile
342          50          55          60
344 tta tta aca agc ctg att tat tca atg gat aag ata gct aac acg gta      240
345 Leu Leu Thr Ser Leu Ile Tyr Ser Met Asp Lys Ile Ala Asn Thr Val
346 65          70          75          80
348 ggg cat tat gaa gct tat agg aaa aaa gag att ttg caa gat aga ttt      288
349 Gly His Tyr Glu Ala Tyr Arg Lys Lys Glu Ile Leu Gln Asp Arg Phe
350          85          90          95
352 att ttt gag ctt att agc cct ata aaa cat gat aaa aat atc atg ata      336
353 Ile Phe Glu Leu Ile Ser Pro Ile Lys His Asp Lys Asn Ile Met Ile
354          100          105          110
356 gag aga aaa aac gct aac gaa ttg gct aaa acc tta aaa ata gac tta      384
357 Glu Arg Lys Asn Ala Asn Glu Leu Ala Lys Thr Leu Lys Ile Asp Leu
358          115          120          125
360 gtc ttt att gat cct cca tac aat tca agg caa tac agc cgg ttt tat      432
361 Val Phe Ile Asp Pro Pro Tyr Asn Ser Arg Gln Tyr Ser Arg Phe Tyr
362          130          135          140
364 cat ctc tat gaa aac cta gtg cag tgg aaa aaa ccc aaa ctc tat gga      480
365 His Leu Tyr Glu Asn Leu Val Gln Trp Lys Lys Pro Lys Leu Tyr Gly
366 145          150          155          160
368 aca gct tta aag cca tca tgc gag aac atg agc gaa tat tgc cgc tct      528
369 Thr Ala Leu Lys Pro Ser Cys Glu Asn Met Ser Glu Tyr Cys Arg Ser
370          165          170          175
372 aat gcc aag aaa gaa ttg agc gat tta att gaa aaa cta gat tgt aaa      576

```

VERIFICATION SUMMARY

DATE: 07/30/2004

PATENT APPLICATION: US/10/796,669A

TIME: 11:26:56

Input Set : A:\741179.txt

Output Set: N:\CRF4\07302004\J796669A.raw